PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Use of the nominal group technique to identify UK stakeholder views of the measures and domains used in the assessment of therapeutic exercise adherence for patients with musculoskeletal disorders
AUTHORS	Mallett, R; Mclean, Sionnadh; Holden, Melanie; Potia, Tanzila; Gee, Melanie; Haywood, Kirstie

VERSION 1 – REVIEW

REVIEWER	Sarah Dean University of Exeter, UK
REVIEW RETURNED	19-Jun-2019

GENERAL COMMENTS	Use of the nominal group technique to identify stakeholder views of the measures and domains used in the assessment of therapeutic exercise adherence. BMJ-Open – 2019 – 031591 Thank you for asking me to review this interesting and useful paper. It has been well written and presented and helps towards addressing a gap in the research regarding the need and development of robust and meaningful measures of adherence to therapeutic exercise for people with musculoskeletal disorders. My initial more detailed comments are some suggestions for refinements to your manuscript and secondly there are some minor corrections to be considered for improving clarity or consistency. Page 6. My first main concern is the section on Patient and Public Involvement (PPI). I do not think the way patients contributed to this study can be considered PPI in the true sense of how this activity is defined: for your information the following is copied from the Involve website: What public involvement in research is not Researchers and others use different words to describe public involvement, for example words such as engagement and participation. In this publication when we use the term 'public involvement' we are not referring to researchers raising awareness of research, sharing knowledge or engaging and creating a dialogue with the public. We are also not referring to the recruitment of

	patients or members of the public as participants in research. https://www.invo.org.uk/posttyperesource/what-is-public-involvement-in-research/. Accessed 19/06/2019. Thus PPI work is not the same as including patients as research participants in stakeholder workshops – in your study were these people giving consent and effectively 'signed up' as research participants under the auspices of your ethical approval? (I was not entirely sure – which suggests some clarification is needed), if this is the case then this is not PPI. Nor does contributing to the voting items after the NGT event constitute true PPI, this is all part of the research participation by both patients and health professionals. Patients certainly contributed as research participants in probably a very equal way as afforded to the health professionals – and this is a strength of your work, but I think this section needs to be amended to ensure you are not claiming any PPI had occurred (at the moment you just state there was no public involvement in study design). Page 9, line 3. I suggest indicating what some of these omissions are, for example the measure published by Newman-Beinart et al 2016 seems a very relevant, and well developed, adherence measure for people doing exercise for low back pain. This was published before your review (McLean et al, 2017) but I could accept that it might not have been in time for it to be included in your review but perhaps some comment about why it was not used in the plan for the NGT would be useful; or instead indicate in this article that you are now aware of such developments - as this would enhance the currency of this article. Page 9, last paragraph. You claim to have identified the lack of reliability, validity and acceptability of measures yet I see no data presented on the two former criteria (I do not disagree with the point that these things are lacking – apart from perhaps the Newman-Beinart et al 2016 measure – but I do not think this article set out to examine reliability or validity). Either amend this sentence or present a summary of the psychometrics (or gaps in psychometrics) for these measures. Page 10. You make the point that the first step in development of a future measure is to establish a conceptual model / theoretical framework. Again, I do not disagree with this premise but believe much of this theoretical work has already been done and I would be concerned about spending a lot of resource developing a new theory or model when there are plenty already in existence – it may be they need adapting to your specific area of musculoskeletal exercise adherence but even that may have already been done (see for example Newman-Beinart PhD thesis that underpinned the 2016 measure). There are also other areas of treatment adherence research and theory development that are very well advanced (e.g. Horne et al, 2019), so again rather than start again I

	would recommend the next steps are very much about adapting and refining what is already known about measuring treatment adherence. Minor points Title: I suggest adding 'in musculoskeletal disorders' – as you have done in article summary. Page 3, last line. I suggest adding the point that non-adherence is also a 'missed opportunity for therapeutic benefit' for patients. Page 5, line 10. Is it 'deliberate'? Page 5. Not sure what happened to stage 3. Page 6 line 30. Does SRQR need to be in full (I couldn't see where it had been mentioned before but may have missed this). Page 6 line 56. 'Data' is a plural word, so should be 'data were'. Page 8 line 35-41. Seems a bit repetitive and could be edited to be more concise. Page 8 line 51 – I do not think they were research partners, at least not from my reading of your manuscript (hence comments about PPI). Maybe I have missed something and my concern about PPI is unfounded, and they were your partners throughout. I apologies if I have got this wrong, editing your manuscript to really clarify whether this was work carried out in partnership with patients and public (i.e. true PPI) or whether you were using patients as research participants (who were afforded equal standing to the health professional participants) seems an important refinement to make to the manuscript. Page 9 line 19, yes I agree! Page 9 line 49, yes I agree this is needed (a valid and reliable measure) but this is not what your work has been about (validity and reliability). Page 10 line 11. I agree that quantification and determinants are a useful way to construct scales. Page 10 line 19, suggest this has highlighted the 'known' complex nature of exercise adherence; I do not think this complexity is newly identified by your work. Page 10 line 22. Check if you are using the term 'therapist' or 'clinician'. Table 2 – says 14 items but there are 12 shaded and 16 in total…….can you clarify?

	Table 2. Items 1 and 7 – these are asking about two things 'amount' and 'frequency', and this might need careful consideration and differentiated for any future work. Table 2 similarly items 13 and 14 are for both physical and mental demands – differentiating out these demands would be useful in terms of clinical utility (and how you might then intervene to support someone's adherence). Table 3 regarding the two items related to resuming exercise after a forced break, these did not meet your threshold for inclusion yet they merit some further comment – relapse management for people is a really important part of therapeutic exercise adherence especially for those with long term conditions. I can accept that your plans for a measure might not include this domain but it is an area that needs consideration. Summary I have thoroughly enjoyed reading your paper and the work you have undertaken, you are definitely on the right lines for progressing this topic of research, I only urge you not to take up precious time and resource re-inventing in physical therapy what has already been happening in psychology / health psychology and sports psychology in the areas of exercise adherence theory and measurement development. I wish you well with your research endeavours. References Newman-Beinart, N. A., Norton, S., Dowling, D., Gavriloff, D., Vari, C., Weinman, J. A., & Godfrey, E. L. (2016). The development and initial psychometric evaluation of a measure assessing adherence to prescribed exercise: The Exercise Adherence Rating Scale (EARS). *PHYSIOTHERAPY*, *102*(4). https://doi.org/10.1016/j.physio.2016.11.001 Horne, R., Cooper, V., Wileman, V., Chan, A. Supporting Adherence to Medicines for Long-Term Conditions: A Perceptions and Practicalities Approach Based on an Extended Common-Sense Model. European Psychologist (2019), 24, pp. 82-96. https://doi.org/10.1027/1016-9040/a000353.

REVIEWER	Dr Emma Godfrey King's College London
REVIEW RETURNED	12-Jul-2019

GENERAL COMMENTS	This paper addresses an important area that requires further research, namely how to assess exercise adherence. The authors used a nominal group technique (NGT) to evaluate six current exercise adherence measures, which they had identified from their recent systematic review published in 2017. This seems a good first step but other stake holders' opinions are needed, as the authors acknowledged in their discussion. The problem with their

	approach was that they did not do an up-to-date literature search before starting this study and therefore unfortunately missed a suitable measure developed by another research group. The Exercise Adherence Rating Scale (EARS) {Newman-Beinart, 2017} was recently developed and should have been included in this study. Construct validity has been assesed and an exploratory factor analysis showed good internal consistency (α=.810) and reliability {Newman-Beinart, 2017}. This is currently the only standardised, validated measure to assess self-reported exercise adherence and it has also been assessed for comprehansion and face validity in people with persistent MSK pain (Meade et al. 2018). Its popularity and use continues to grow worldwide, as it is currently being translated into twelve languages, therefore it is a great pity that this measure was omitted from this process. The conclusions of the arcticle are therefore doubtful as although all six measures were rejected due to a perceived lack of suitability for routine practice or clinical trials, the EARS could be suitable and indeed was designed for that purpose. It is helpful that 25 domains of exercise adherence have been indentified but further research to develop a measure of adherence with therapeutic exercise may not be necessary.

VERSION 1 – AUTHOR RESPONSE

Dear Reviewers, thank you for your comments. These have been very helpful in improving the accuracy and the currency of our manuscript. We have been able to take on board all your comments, please see our responses to individual comments below.

Reviewer 1	
c1	Page 6. My first main concern is the section on Patient and Public Involvement (PPI). I do not think the way patients contributed to this study can be considered PPI in the true sense of how this activity is defined: for your information the following is copied from the Involve website: What public involvement in research is not Researchers and others use different words to describe public involvement, for example words such as engagement and participation. In this publication when we use the term 'public involvement' we are not referring to researchers raising awareness of research, sharing knowledge or engaging and creating a dialogue with the public. We are also not referring to the recruitment of patients or members of the public as participants in research. https://www.invo.org.uk/posttyperesource/what-is-public-involvement-in-research/. Accessed 19/06/2019. Thus PPI work is not the same as including patients as research participants in stakeholder workshops – in your study were these people giving consent and effectively 'signed up' as research participants under the auspices of your ethical approval? (I was not entirely sure – which suggests some clarification is needed), if this is the case then this is not PPI. Nor does contributing to the voting items after the NGT event constitute true PPI, this is all part of the research participation by both patients and health professionals. Patients certainly contributed as research participants in

	probably a very equal way as afforded to the health professionals – and this is a strength of your work, but I think this section needs to be amended to ensure you are not claiming any PPI had occurred (at the moment you just state there was no public involvement in study design).
R	Thankyou for pointing out our error. As you say the PPI section in our manuscript does not do justice to the level of PPI that did take place. As a multi stage project with a team approach not all members were involved in the preparatory PPI and methodology design stages. We have reviewed these stages and significantly reworked this section to show the PPI involvement as it was. In the manuscript word count has meant that we have kept this section fairly short. However, a fuller explanation of our PPI activities is provided below. The opinions of patients, and of physiotherapists, have directly informed the development of this research proposal, through the views and experiences that they shared within a previous study at Keele University (The ABC-Knee study), and through workshops completed with PPI and clinical groups within South Yorkshire Collaborative Local Research Network. The ABC-Knee study (attitudes and beliefs concerning knee pain) highlighted that patients recognised the importance of exercise adherence, felt it was their own responsibility to maintain therapeutic exercise programmes, but identified many barriers to maintaining exercise and activity in the presence of musculoskeletal pain. Physiotherapists also recognised the importance of adherence in determining outcomes from exercise programmes, but overall there was a lack of robust measurement of exercise adherence within physiotherapy practice, which could inhibit the use of adherence enhancing interventions with patients who need them the most. This highlighted the clear need to identify the best available measures of exercise adherence, not only for research purposes, but also for use in physiotherapy clinical practice. The proposed aims, design and methods of this research proposal have been discussed with the Barnsley Consumer Research Advisory Group (CRAG) and the AHP special interest group; these are respectively the PPI and the Allied Health Professionals special interest groups of the South Yorkshire Collaborative Local Research Network (SY CLRN). Both groups agreed that the project will be valuable because of the scale of non-adherence to exercise for musculoskeletal disorders within physiotherapy practice. A number of views of the Barnsley CRAG group in particular have shaped the development of this project as follows: • The Barnsley CRAG thought it was important to have good transport links to facilitate the attendance of patient representatives at the consensus group, therefore we planned to conduct the workshop at the city centre campus of Sheffield Hallam University which has extremely good bus, rail and road transport connections close at hand. • They thought that patient representatives may feel a little daunted by the prospect of the workshop and suggested a range of organisational considerations and activities that would help e.g. having a slightly increased ratio of patients; a friendly engaging facilitator, warm up tasks to help involve the patient representatives and get them chatting with other groups before the workshop related tasks begin. • The Barnsley CRAG supported the use of the Nominal Group Technique because they felt that it would offer all individuals the chance to contribute equally through voting, even if not all participants wanted to speak and it would have the effect of equalising the contributions through

	the workshop.
c2	Page 9, line 3. I suggest indicating what some of these omissions are, for example the measure published by Newman-Beinart et al 2016 seems a very relevant, and well developed, adherence measure for people doing exercise for low back pain. This was published before your review (McLean et al, 2017) but I could accept that it might not have been in time for it to be included in your review but perhaps some comment about why it was not used in the plan for the NGT would be useful; or instead indicate in this article that you are now aware of such developments - as this would enhance the currency of this article.
R	We agree that enhancing the currency of this article is an appropriate thing to do. We have acknowledged in the line suggested and several other places in the discussion section the development of the EARS and some of the similarities between our work and the doctoral work of Dr Newman-Beinart.
C3	Page 9, last paragraph. You claim to have identified the lack of reliability, validity and acceptability of measures yet I see no data presented on the two former criteria (I do not disagree with the point that these things are lacking – apart from perhaps the Newman- Beinart et al 2016 measure – but I do not think this article set out to examine reliability or validity). Either amend this sentence or present a summary of the psychometrics (or gaps in psychometrics) for these measures.
R	Yes, agreed. We have removed the wording reliability and validity in this section
c4	Page 10. You make the point that the first step in development of a future measure is to establish a conceptual model / theoretical framework. Again, I do not disagree with this premise but believe much of this theoretical work has already been done and I would be concerned about spending a lot of resource developing a new theory or model when there are plenty already in existence – it may be they need adapting to your specific area of musculoskeletal exercise adherence but even that may have already been done (see for example Newman-Beinart PhD thesis that underpinned the 2016 measure). There are also other areas of treatment adherence research and theory development that are very well advanced (e.g. Horne et al, 2019), so again rather than start again I would recommend the next steps are very much about adapting and refining what is already known about measuring treatment adherence.
R	Thankyou for these recommendations. We couldn't agree more therefore have amended the statement to suggest possible integration of these theoretical underpinnings.

	Minor points
c5	Title: I suggest adding 'in musculoskeletal disorders' – as you have done in article summary.
R	Amended.
c1	Page 3, last line. I suggest adding the point that non-adherence is also a 'missed opportunity for therapeutic benefit' for patients.
R	We have incorporated this into the first paragraph of the Introduction section
c6	Page 5, line 10. Is it 'deliberate'?
R	For clarity we have removed the word "deliberated"
c7	Page 5. Not sure what happened to stage 3.
R	For clarity we have removed reference to the 3 staged approach.
c8	Page 6 line 30. Does SRQR need to be in full (I couldn't see where it had been mentioned before but may have missed this).
R	You are quite correct. We have added the "Standards for Reporting Qualitative Research" in this sentence.
c9	Page 6 line 56. 'Data' is a plural word, so should be 'data were'.
R	Thank you, corrected.
c10	Page 8 line 35-41. Seems a bit repetitive and could be edited to be more concise.

R	Thankyou, we have edited these lines and reworked the middle part of the first discussion paragraph. Hopefully this is more concise now.
c11	Page 8 line 51 – I do not think they were research partners, at least not from my reading of your manuscript (hence comments about PPI). Maybe I have missed something and my concern about PPI is unfounded, and they were your partners throughout. I apologies if I have got this wrong, editing your manuscript to really clarify whether this was work carried out in partnership with patients and public (i.e. true PPI) or whether you were using patients as research participants (who were afforded equal standing to the health professional participants) seems an important refinement to make to the manuscript.
R	We have amended from "research partners" to "involvement of patients in the research process". We hope that this sufficiently moves this meaning away from PPI which is a different element and a separate process that has been fully explained under the PPI heading in the manuscript.
c12	Page 9 line 19, yes I agree!
R	No response required
c13	Page 9 line 49, yes I agree this is needed (a valid and reliable measure) but this is not what your work has been about (validity and reliability).
R	Thank you we agree. We have amended this sentence.
c14	Page 10 line 11. I agree that quantification and determinants are a useful way to construct scales.
R	No response required
c15	Page 10 line 19, suggest this has highlighted the 'known' complex nature of exercise adherence; I do not think this complexity is newly identified by your work.
R	Yes agreed. We have changed "highlighted" to "confirmed" which hopefully moves away from a

	suggestion that we have newly identified this complexity but supported others' findings.
c16	Page 10 line 22. Check if you are using the term 'therapist' or 'clinician'.
R	Thank you. 'Clinician' has been used consistently to ensure clarity.
c17	Table 2 – says 14 items but there are 12 shaded and 16 in total…….can you clarify?
R	This refers to the number of participants/stakeholders that undertook the voting. It has been removed to aid clarity.
c18	Table 2. Items 1 and 7 – these are asking about two things 'amount' and 'frequency', and this might need careful consideration and differentiated for any future work.
R	Yes, we agree with this
c19	Table 2 similarly items 13 and 14 are for both physical and mental demands – differentiating out these demands would be useful in terms of clinical utility (and how you might then intervene to support someone's adherence).
R	Yes, we agree
c20	Table 3 regarding the two items related to resuming exercise after a forced break, these did not meet your threshold for inclusion yet they merit some further comment – relapse management for people is a really important part of therapeutic exercise adherence especially for those with long term conditions. I can accept that your plans for a measure might not include this domain but it is an area that needs consideration.
R	We do agree with your assessment that resuming exercise after a forced break/relapse management is an important clinical consideration that warrants further consideration. However, we feel that we are not able to do justice to such an important issue within the constraints of the author guidelines and presentation of these specific findings from the consensus voting. We hope to pick up on this important area in future research.
c21	Summary

	I have thoroughly enjoyed reading your paper and the work you have undertaken, you are definitely on the right lines for progressing this topic of research, I only urge you not to take up precious time and resource re-inventing in physical therapy what has already been happening in psychology / health psychology and sports psychology in the areas of exercise adherence theory and measurement development. I wish you well with your research endeavours.
R	Thank you very much for taking the time and trouble to read this so thoroughly and carefully. We appreciate the effort that you have gone to and the advice and recommendations provided here and throughout your comments.
Reviewer 2	
c22	This paper addresses an important area that requires further research, namely how to assess exercise adherence. The authors used a nominal group technique (NGT) to evaluate six current exercise adherence measures, which they had identified from their recent systematic review published in 2017. This seems a good first step but other stake holders' opinions are needed, as the authors acknowledged in their discussion.
R	Thankyou
c23	The problem with their approach was that they did not do an up-to-date literature search before starting this study and therefore unfortunately missed a suitable measure developed by another research group. The Exercise Adherence Rating Scale (EARS) {Newman-Beinart, 2017} was recently developed and should have been included in this study. Construct validity has been assessed and an exploratory factor analysis showed good internal consistency (α=.810) and reliability {Newman-Beinart, 2017}. This is currently the only standardised, validated measure to assess self-reported exercise adherence and it has also been assessed for comprehension and face validity in people with persistent MSK pain (Meade et al. 2018). Its popularity and use continues to grow worldwide, as it is currently being translated into twelve languages, therefore it is a great pity that this measure was omitted from this process.
R	Thank you for this comment. Yes, it is very unfortunate that the EARS measure was not included here. This was because this study was part of a multi-stage project, where the literature review identifying adherence measures was conducted prior to the publication of the EARS measure. We will definitely be sure to include the EARS in any updates of our systematic review of adherence measures.
c24	The conclusions of the article are therefore doubtful as although all six measures were rejected due to a perceived lack of suitability for routine practice or clinical trials, the EARS could be suitable and indeed was designed for that purpose. It is helpful that 25 domains of exercise adherence have been identified but further research to develop a measure of adherence with therapeutic exercise may not be necessary.

R	We do accept that other measures have been developed and have since been published from when our study was complete, written and has undertaken the publication submission and review process. As our findings suggest this was definitely needed with evidence of good measurement properties for the EARS that would have been ideal for comparison against those available at the time. The conclusions of this article are based on the material that was available at the point of extraction and subsequent consideration of the stakeholders assembled. We have made needed clarifications to the conclusions which we hope will be satisfactory to fully acknowledge the potential of subsequently published work and the need for its consideration in future research.

VERSION 2 – REVIEW

REVIEWER	Sarah Dean University of Exeter Medical School, UK
REVIEW RETURNED	04-Nov-2019

GENERAL COMMENTS	The authors have done a very good job with addressing my concerns and re-working the manuscript to better represent the contribution of this work in the context of other recently emerging research which is directly relevant to this topic. I believe some very minor revisions are still required to ensure the reader fully understands the contribution of this work versus what is now also available in the published literature. A further review of these revisions is not necessary if the editor is willing to check these relatively straightforward changes have been made. Page 2 Abstract, conclusion first line: replace 'current' with 'these six' Page 2 Article summary, last line of page: delete 'current'. Page 4 Methods, second line of section: replace 'current' with 'these six' Page 8 Discussion, line 4 of section: insert 'six' - 'agreed that the six identified...' Page 8 Discussion line 7-8 of section rewrite to make clear it is about the measures from the earlier review: 'which concluded that the exercise adherence measures covered in the review are of poor quality....' Page 8 Discussion line 10 of section: replace 'existing' with 'these' Page 8 Discussion line 14 of section: suggest delete 'may be' and replace with 'is' a promising new measure' (but I agree with authors still further work remains to be done - see next point). Page 8 Discussion line 15 of section: edit sentence to include 'remains' and 'further': 'There remains a pressing need to prioritise further development....' Page 9 Discussion, 3rd line on page: suggest insert 'accuracy': '(e.g. the amount, frequency and accuracy of completed exercise...' Page 10 Discussion first line of second paragraph: insert 'these six': 'This study has identified that these six measures of....'

REVIEWER	Dr Emma Godfrey King's College London

REVIEW RETURNED	19-Nov-2019

GENERAL COMMENTS	This is a much improved manuscript, as you have included the recent measure of exercise adherence, the EARS, and discussed it appropriately. Perhaps the exclusion of this and subsequent limitation of your work could be reflected more through out your write up, but otherwise this is now well written and complete and tackles an important area.

VERSION 2 – AUTHOR RESPONSE

I thank you and the reviewers again for your time and response. We accept the comments from both reviews and agree the suggestions further acknowledge more contemporary research exists, especially the need to remove the phrase 'current measures'. I have have made the revisions as suggested to ensure readers are aware our work's contribution to the body of evidence in light of subsequently published literature. These changes throughout the script have repeatedly emphasised the capturing of measures and stakeholders view's at that point in time with explicit recognition of exclusion of the EARS which is a study limitation leading to the conclusions of the remaining need for further work in this area.